NBER WORKING PAPER SERIES

FEDERAL LIFE SCIENCES FUNDING AND UNIVERSITY R&D

Margaret E. Blume-Kohout
Krishna B. Kumar
Neeraj Sood

Working Paper 15146
http://www.nber.org/papers/w15146

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2009

We are very grateful to Bhaven Sampat and Pierre Azoulay for providing analytic data files, and to Sean Nicholson and RAND Labor & Population seminar attendees for helpful comments. M. Blume-Kohout also gratefully acknowledges partial funding from the Ewing Marion Kauffman Foundation, and the Kip & Mary Ann Hagopian and Anne & James Rothenberg dissertation awards at the Pardee RAND Graduate School. These sponsors have had no role in our study design, analysis, interpretation of data, or writing, and will not influence the eventual decision whether to submit this paper for publication. An earlier draft of this paper was also submitted as part of a workshop organized and funded by the Office of Science Policy Analysis, Office of the Director, National Institutes of Health. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Federal Life Sciences Funding and University R&D
Margaret E. Blume-Kohout, Krishna B. Kumar, and Neeraj Sood
NBER Working Paper No. 15146
July 2009
JEL No. H5,I1,I23,O3

ABSTRACT

This paper investigates the impact of federal extramural research funding on total expenditures for life sciences research and development (R&D) at U.S. universities, to determine whether federal R&D funding spurs funding from non-federal (private and state/local government) sources. We use a fixed effects instrumental variable approach to estimate the causal effect of federal funding on non-federal funding. Our results indicate that a dollar increase in federal funding leads to a $0.33 increase in non-federal funding at U.S. universities. Our evidence also suggests that successful applications for federal funding may be interpreted by non-federal funders as a signal of recipient quality: for example, non-PhD-granting universities, lower ranked universities and those that have historically received less funding experience greater increases in non-federal funding per federal dollar received.

Margaret E. Blume-Kohout
RAND Corporation
1776 Main Street
Santa Monica, CA 90407
and
University of Waterloo
Waterloo, Ontario, Canada
mblumek@artsservices.uwaterloo.ca

Krishna B. Kumar
RAND Corporation
1776 Main Street
Santa Monica, CA 90407
kumar@rand.org

Neeraj Sood
RAND Corporation
1776 Main Street
Santa Monica, CA 90407
and NBER
neeraj_sood@rand.org

1. Introduction

Federal agencies spend billions of taxpayer dollars funding academic research each year. Funding for academic research, especially in the life sciences, has risen rapidly over the last decade. In 2007, total federal obligations for research and development (R&D) at academic institutions totaled over $25 billion, of which $15.5 billion was provided by the National Institutes of Health (NIH), primarily to support basic and applied research in the biomedical sciences.[1] The American Recovery and Reinvestment Act of 2009 (ARRA) provided an additional $8.2 billion to NIH to fund extramural life sciences R&D. The oft-cited justification for this level of public expenditure is that publicly-funded biomedical research results in better medical treatments and even preemption of disease (National Institutes of Health, 2008). Under ARRA, increases in federal funding for biomedical research are further justified as a means to speed economic recovery, by increasing overall spending and employment opportunities. However, despite these laudable aims, economic theory suggests that federal funding for R&D could "crowd out" investment by fully-informed non-federal funders, thereby reducing its effectiveness.

Increased federal investment could also cause individual investigators at universities to substitute among funding sources. Because applying for funding is a costly activity, and because increases in income may reduce the perceived utility of additional funds, an increase in universities' federal funding might make them less inclined to pursue other funding sources. For either or both of these reasons, a dollar increase in federal funding could yield less than a dollar increase in total R&D.

On the other hand, private and public funding could be complements instead of substitutes. For example, by subsidizing investment in capital or equipment, or by supporting

skill development, federal funding might increase the productivity of a university, making it more competitive in the market for non-federal research funding. If funding sources are complementary, a dollar increase in federal funding would increase total R&D by more than a dollar.

Even if federal funding does not directly improve productivity, it may serve as a signal of university quality. For example, due to the extensive peer review conducted by the NIH and other federal agencies, non-federal funders seeking university partners may view successful applications for federal funding as a signal of university quality. One might therefore expect a stronger signaling effect at universities with less established research reputations, about which potential non-federal investors have less information. In this case, a dollar increase in federal funding should again yield more than a dollar increase in total funding for R&D, but with larger effects at universities with smaller research portfolios and less established reputations.

Finally, it is possible that federal dollars may be the sole source of financial support for some types of R&D, and thus neither substitute for nor complement private R&D funding. For example, some have argued basic science is a public good, and would therefore be underfunded by private sources. If this were the only effect operating, in the short run one would expect changes in federal funding to have no impact on non-federal funding: that is, a dollar increase in federal funding would simply increase total R&D funding by one dollar.

In this paper, we estimate the causal effect of federal funding on non-federal funding for life sciences R&D at U.S. universities and colleges. Our empirical models account for potential bias due to university characteristics that could be associated with both higher federal and higher non-federal funding, such as the number and quality of faculty, supportiveness of the university research environment, and the institution's reputation.

This question of whether federal funding substitutes for or complements other funding has been previously studied in the literature. For example, David et al. (2000) reviewed literature on government-issued contracts to the private sector and policies such as R&D tax credits, and found mixed results: some studies suggested public R&D complements private R&D, while others found that public R&D crowds out private R&D. Our research differs from the studies reviewed by David et al. (2000) in that we focus on university recipients, and the extent to which public (federal) R&D funding impacts R&D funding received from all non-federal sources. However, despite these differences, the warnings in David et al. (2000) about latent omitted variables and selection bias are relevant to our work. Diamond (1999) used aggregate, annual time series data from 1953-1995 to study the connection between federal and private spending on basic research in science, and found that a dollar increase in federal spending was associated with an additional $0.08 in academic spending. Like David et al. (2000), Diamond (1999) cautions the reader about omitted latent variables, but he does not apply any econometric correction.

Some more recent studies have used instrumental variables and regression discontinuity designs to address the omitted variables issue. For example, Payne (2001) and Payne and Siow (2003) use instrumental variables to investigate the possible causal relationships between federal funding and (a) philanthropic funding, and (b) research outcomes, including patents and publications. Payne (2001) finds, when restricting observations to years post-1980, that a dollar increase in federal research funding increases private donations by $0.64 to $0.68, with no significant difference between private and public universities. Payne and Siow (2003) find that a $1 million (1996 dollars) increase in federal funding to a university yields 10 more publications and 0.2 patents. In contrast, using a regression discontinuity design, Jacob and Lefgren (2007)

find only modest increases in publication productivity for recipients of NIH R01 research project grants, which they suggest may be due to substitution: on the margin, loss of an NIH grant simply causes researchers to shift to another source of funding.

Our work complements prior research by exploring a mechanism through which federal biomedical research funding might influence eventual commercialization (i.e., by attracting complementary funding from private sources), and extends Payne (2001) to consider other non-federal funding sources besides philanthropic donations. In addition, we construct a new instrument that allows us to infer causality despite the presence of unobserved variables that may impact both federal and non-federal funding at a given university over time. Finally, our paper contributes to the existing literature by examining the particular case of federal life sciences funding (mainly originating from the NIH), highlighting heterogeneous effects across universities with differing characteristics, and extending earlier work with more current data from a variety of sources.

To address potential omitted variable bias, we estimate an instrumental variable (IV) model with university fixed effects, using predicted NIH funding as an instrument for federal life sciences funding. To generate this instrumental variable, we first calculate the share of each university's R&D funding awarded by each NIH Institute or Center (NIC) in the base year of our analysis. Because each of the NICs specializes in particular diseases, areas of human development, or aspects of research support (see Smith (2006)), each university's distribution of base-year funding across NICs reflects its particular research specialization. We then predict the total NIH funding each university receives in subsequent years based on year-to-year changes in the budgets of the individual NICs, assuming that the university's relative specialization, i.e., the share of funding it receives from each NIC, remains constant throughout the study period. We

use this predicted NIH funding variable as an instrument for actual funding in our empirical analysis.

The challenge in implementing the IV estimator is to find an instrument that is strongly correlated with changes in federal research funding at universities, but is uncorrelated with other types of shocks that might affect non-federal funding. Because year-to-year appropriations for the different NICs are determined via political processes, the predicted funding levels described above are unlikely to be related to changes in other factors that drive a particular university's research capabilities and, therefore, its likelihood of receiving non-federal funding. We present results from several tests of this assumption. For example, we find no correlation between year-to-year variation in NIC budgets and contemporaneous variation in related industry R&D. We also find no relationship between the share of funding universities received from each NIC and other observable characteristics associated with higher non-federal R&D. The results we report here are also robust to changes in calculation of predicted NIH funding, such as using an earlier base-year to calculate NIC shares. Predicted NIH funding is therefore a useful instrument for the total federal life sciences R&D funding a university receives, and allows us to attribute causality to federal life sciences and NIH funding.

We find that an additional dollar of federal funding for life sciences research increases non-federal R&D funding by $0.33. The point elasticity (at the average) for non-federal funding to lagged federal funding is 0.51. Our results also indicate that non-PhD-granting universities and those that have historically received less funding from all sources may experience greater increases in non-federal funding for each federal dollar. This suggests successful applications for federal funding may be interpreted by non-federal funders as a signal of university quality.

2. Data and Identification Strategy

2.1 Data Sources

Data for this paper are derived from two sources: the National Science Foundation (NSF) Survey of Research and Development Expenditures at Universities and Colleges, and administrative records maintained by the Office of Extramural Research at NIH.

The NSF Survey population includes institutions granting bachelors or higher degrees in science and engineering (S&E) fields, and spending at least $150,000 annually in S&E research and development (R&D). Surveyed institutions report their S&E R&D expenditures by funding source and field. For example, these data include the amount of life sciences R&D funding received from federal versus non-federal (industry, state and local government, institutional, or other) sources. For this analysis, we extracted universities' total and federally-funded R&D expenditures by year and field for 1998 through 2006. Our dependent variable, non-federal life sciences R&D funding, was calculated by subtracting federally-funded life sciences R&D expenditures from total life sciences R&D expenditures. Our key explanatory variable is federally-funded life sciences R&D expenditures, lagged by one year, and is also derived from this survey. Life sciences R&D includes research in agricultural, biological, and medical sciences, as well as allied health professions; however, since 1998 over half of academic R&D expenditures in the life sciences have been for medical research, and this share has continued to grow over time. Finally, for some models we include covariates for universities' federally- and non-federally funded R&D in fields other than life sciences. To generate these covariates, we subtract life sciences R&D from total R&D expenditures to calculate non-life-sciences R&D expenditures for each institution-year. Our dataset also includes other institutional

characteristics, such as whether the institution is public or private, and whether it grants PhDs in S&E fields. These additional characteristics allow us to explore possible signaling effects.

NIH administrative data for each grant and contract awarded in fiscal years 1997 through 2006 include the grant or contract unique ID number, the fiscal year of the award, principal investigator's institution (including institution name, city, and state), and the financial amount of the award.[2] As discussed in section 2.2, we use NIH award data from 1997 to calculate universities' base-year funding shares by NIC. We also use subsequent years' data in a regression predicting non-federal funding as a function of actual NIH awards. This measure differs from our key explanatory variable in that, while NIH is the lead federal agency funding academic life sciences research, universities may also receive life sciences R&D funding from other federal agencies, such as the Centers for Disease Control (CDC) or Food and Drug Administration (FDA). In addition, although most NIH extramural R&D funding supports basic and applied life sciences research, NIH also funds research in other fields. Thus, though actual NIH funding is strongly correlated with universities' total life sciences R&D expenditures, for consistency we rely on the NSF Survey data for our dependent and key explanatory variables.

We matched institutions across these two datasets in an iterative process. First, we found all exact matches by institution name and state. Then, we extracted all remaining awardees in the NIH data that were coded as institutions of higher education, and matched these institutions by hand with those listed in the NSF survey. Finally, we included in our analytic dataset only those institutions for which NSF survey data were available for each year in our study period, 1998-2006, and for which NIH awards were observed in the administrative data in 1997, which limited our dataset to 272 institutions.

Table 1 provides descriptive statistics for our analytic dataset. Approximately one-third of the 272 institutions in our panel are private, and over three-quarters grant PhDs in S&E fields. On average, about 60% of university R&D expenditures were funded by the federal government (62% for life sciences, and 60% for other fields).

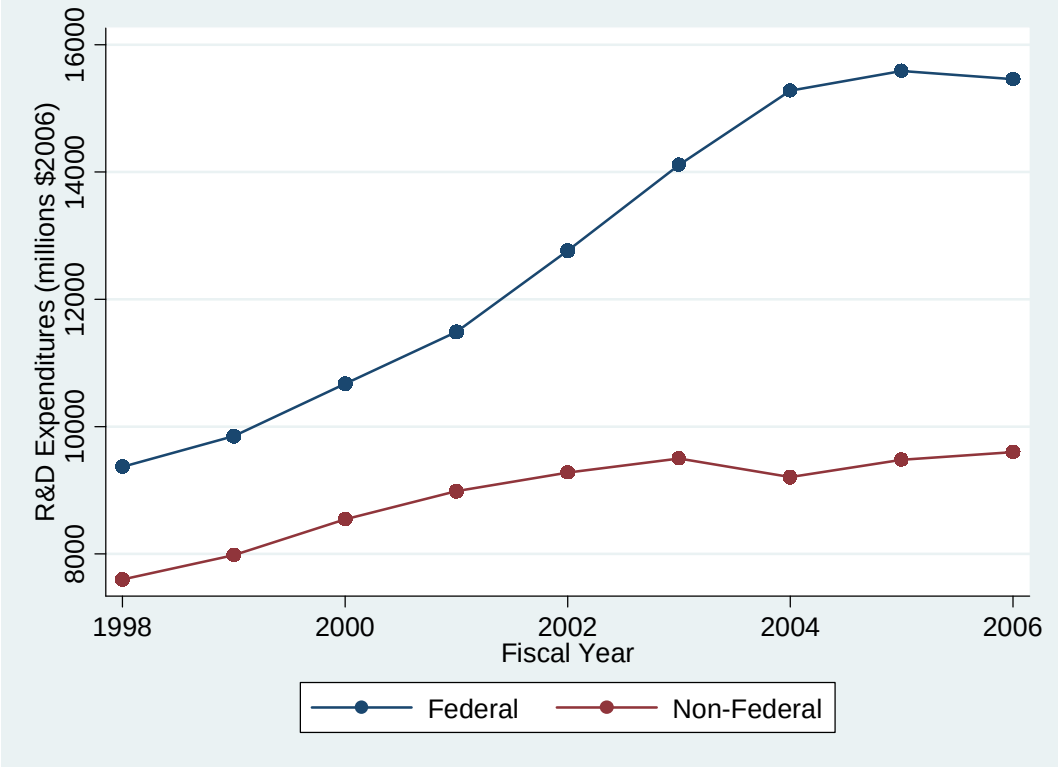
Figure 1 shows the differences in growth for federal and non-federal life sciences R&D funding for our panel. From 1998 through 2001, university life sciences R&D funding from both sources grew at about the same rate, but in subsequent years, federal funding outpaced non-federal funding. In 2006, the universities in our panel spent over $25 billion on life sciences R&D, representing approximately 87% of the national total.[3]

Table 1: Descriptive statistics for university R&D expenditure data

Variable	Mean	Standard Deviation
Federal Funding for Life Sciences	46.8	74.8
Non-Federal Funding for Life Sciences	28.5	44.4
Federal Funding for Other Fields	31.2	51.0
Non-Federal Funding for Other Fields	21.1	33.1
PhD-Granting Institutions	77.9%	
Private Institutions	31.6%	
Number of Institutions	272	

Note: Author's calculations based on data from National Science Foundation (NSF) Survey of Research and Development Expenditures at Universities and Colleges. All amounts reported in constant 2006 dollars, inflated using the Biomedical Research and Development Price Index (BRDPI). Percent of institutions granting PhDs in S&E fields based on institution's highest degree granted in 1997. Reported standard deviations are calculated between panel institutions.

Figure 1: R&D expenditures by funding source, panel of 272 universities, 1998-2006



2.2 Empirical Methods

We employ several empirical strategies to estimate the relationship between federal and non-federal R&D funding at universities. For all analyses, the unit of observation is the university-year, and standard errors are clustered at the university level to accommodate serial correlation. We first conduct descriptive analyses to examine the association between federal and non-federal funding for life sciences at research universities. Then, we estimate four different sets of multivariate linear regression models to investigate the possibility of a causal relationship between federal and non-federal R&D funding. Each set of multivariate regression models incrementally controls for observed and unobserved university characteristics that could bias our estimates of the causal relationship between federal and non-federal funding.

Our first regression estimates the simplest model:

$$NonFederal_{u,t} = \alpha_0 + \alpha_1 Federal_{u,t-1} + \tau + \varepsilon_{u,t} \quad (1)$$

where $Federal_{u,t-1}$ is the real, lagged federal funding for life sciences R&D at university u in year t - 1; $NonFederal_{u,t}$ is the real non-federal funding for life sciences at university u in year t, τ is a vector of year fixed effects that non-parametrically controls for secular changes in non-federal funding over time (e.g., due to changes in economic conditions); α_0 is a constant; and $\varepsilon_{u,t}$ is the error term. The key coefficient of interest is α_1, which estimates the change in non-federal funding associated with a dollar increase in federal funding the previous year. We have two reasons for using lagged federal funding as our key independent variable. First, as discussed in Section 1, we anticipate that any true crowd-in (or crowd-out) effect should be due to information about changes in federal funding that non-federal funders observe. Because non-federal funders cannot observe concurrent public funding decisions, we lag all federal funding by one year. Second, lagging federal funding insulates us, to a certain extent, from exogenous

shocks that might increase both federal and non-federal funding (e.g., increased interest in funding for a particular disease in which the university has specialized, university hiring of new senior faculty, etc.) in a given year.

However, as discussed in Section 1, α_1 may nonetheless suffer from omitted variables bias. For example, universities with more faculty members or stronger research reputations may attract both greater federal and greater non-federal funding. These omitted variables would yield a positive association between federal and non-federal funding, even if no causal relationship exists. Thus our estimate of α_1 is likely to be biased upward due to this spurious correlation, because the model does not control for university characteristics. Our second regression model addresses this source of bias. We estimate:

$$NonFederal_{u,t} = \beta_0 + \beta_1 Federal_{u,t-1} + \tau + \mu + \varepsilon_{u,t} \quad (2)$$

where μ is a vector of university fixed effects that control for all time-invariant differences across universities. In contrast to equation (1) which exploited variation in funding levels both across and within universities, this model exploits only the variation within individual universities' prior year federal funding to estimate the effect on non-federal funding. In effect, this model estimates whether a university that received more federal funding (relative to its average over the study period) in year t received increased (or decreased) non-federal funding in year t+1. In this specification, β_1 could be biased if time-varying university characteristics are correlated with growth (or decline) in federal and non-federal funding. For example, as discussed by Lawler (2003), growth in non-federal (i.e., industry) funding at top research universities such as the University of California at Berkeley and MIT may reflect strategic initiatives by university administration to diversify funding sources. If increased federal fundraising activity occurred at

the same time, we might erroneously conclude that federal funding caused growth in non-federal funding.

Our third regression model is designed to control for additional bias arising from such time-varying university characteristics, by including the amounts of federal and non-federal funding the university received for other S&E fields:

$$NonFederal_{u,t} = \gamma_0 + \gamma_1 Federal_{u,t-1} + \tau + \mu + \gamma_2 OtherFed_{u,t-1} + \gamma_3 OtherNonFed_{u,t} + \varepsilon_{u,t} \quad (3)$$

In this model, we include lagged federal funding for other fields because overall university reputation (not just within biomedical sciences) as captured by total federal funding may influence non-federal funding decisions, particularly for complementary research. Inclusion of same-year non-federal funding for other fields allows us to control for year-to-year differences in the extent to which universities seek non-federal funding. Because university policy with respect to non-federal funding is likely to be correlated with university quality, excluding these effects could likewise yield biased estimates of γ_1.

Finally, we use instrumental variables (IV) estimation to account for other observed and unobserved time-varying university characteristics that might bias our coefficients. Our instrument for federal life sciences R&D funding is predicted NIH funding. Predicted NIH funding is calculated based on the share of a given university's funding received from each NIC in our base year, and the overall growth (or decline) in each NIC's budget each year. Specifically, predicted NIH funding is given by:

$$\overline{NIH}_{u,t} = NIH_{u,b} \sum_i \left(\frac{Budget_{i,t}}{Budget_{i,b}} \right) * share_{i,u,b} \quad (4)$$

where $NIH_{u,b}$ is the actual NIH funding for university u in our base year, 1997; $Budget_{i,t}$ is the total annual budget for NIH institute i in year t; and $share_{i,u,b}$ is the share of total funding from NIC i for university u in base year b. Equation (4) shows that growth in a given university's predicted NIH

funding is equal to the weighted average growth across NIC budgets, where weights reflect the university's share of funding from each NIC in the base year. We then estimate the following IV model using two-stage least squares:

$$Federal_{u,t} = \theta_0 + \theta_1 \overline{NIH}_{u,t} + \theta_2 OtherFeder\,al_{u,t} + \theta_3 OtherNonFe\,deral_{u,t} + \tau + \mu + \varsigma_{u,t} \quad (5)$$

$$NonFederal_{u,t} = \delta_0 + \delta_1 Federal_{u,t-1} + \delta_2 OtherFederal_{u,t-1} + \delta_3 OtherNonFederal_{u,t} + \tau + \mu + \varepsilon_{u.t} \quad (6)$$

Our decision to analyze the effects of absolute rather than percent changes (i.e., the decision not to log-transform the dependent and independent financial amounts) was based on three observations. First, the large number of institution-years with no federal and/or non-federal R&D funding for S&E fields would require us to drop nearly a hundred universities from our panel (decreasing the number from 272 to 179). Ad-hoc methods to deal with this problem, such as replacing all zero values with $1 or $1000, yield results which strongly depend on the choice of replacement value. Second, exclusion of institutions with historically low levels of funding from either source is counterproductive because we are particularly interested in evaluating the possibility that successful applications for federal funding provide a signal of quality to non-federal funders. The 179 institutions that do receive funding in all years are overwhelmingly PhD-granting (93% vs. 78% in the full panel), and are somewhat more often public than private (72% vs. 68% in the full panel). As one would expect, they also have higher average federal and non-federal funding for both life- and non-life sciences, thus are unlikely to permit an examination of signaling effects. Finally, to the extent that variance in the errors increases with funding levels, heteroskedasticity in the linear model can be accommodated simply by our calculation of robust standard errors.[4,5]

3. Results

3.1 Full Sample Effects

Table 2 shows the results from our multivariate regressions. The dependent variable in each case is non-federal life sciences R&D funding, and the key independent variable is federal life sciences R&D funding. The results from model 1 (corresponding to equation (1)) show that a dollar increase in federal life sciences funding is associated with a $0.47 increase ($p<.001$) in non-federal life sciences funding. However, as discussed above, the lack of any controls for university characteristics in this model likely bias this estimate upwards. For example, universities with larger faculties and/or reputations for higher quality research might receive both more federal and more non-federal funding for life sciences R&D. To address this concern, subsequent models include university fixed effects to control for time invariant university characteristics. Results from model 2 (corresponding to equation (2)) show that after controlling for time invariant university characteristics and secular time trends, a dollar increase in federal funding for life sciences R&D is associated with a $0.31 increase ($p<.001$) in non-federal funding for life sciences R&D. Model 3 (corresponding to equation (3)) adds covariates for federal and non-federal R&D funding in other fields, to control for unobserved time-varying university characteristics. For example, as discussed in section 1, during our sample period a given university's administration might have pushed for diversification in the university's funding portfolio, or a smaller university might have transitioned from a more teaching-focused to a more research-focused institution. Results from this model indicate that a dollar increase in federal funding for life sciences R&D is associated with a $0.29 increase ($p<.001$) in non-federal life sciences R&D funding. As one would expect, in this model we find no significant impact of federal R&D funding for other fields (i.e., non-life-sciences) on non-federal life sciences R&D.

Model 4 (corresponding to equations (4) through (6)) presents the results from our instrumental variables approach, showing that a dollar increase in federal funding leads to a \$0.33 increase ($p<0.001$) in non-federal life science R&D funding. The Durbin-Wu-Hausman null hypothesis states that an ordinary least squares (OLS) estimator of the same equation would yield consistent estimates; our results soundly reject consistency of the OLS estimator ($p<0.01$). Our results also show that predicted NIH funding is a strong predictor of federal life sciences funding, with a F-statistic for the first-stage regression of 244.32.

Table 2: Effect of federal R&D funding on non-federal R&D funding for life sciences at U.S. universities

	(1)	(2)	(3)	(4)
Federal Funding for Life Sciences	0.47	0.31	0.29	0.33
	[9.18]***	[5.91]***	[5.98]***	[4.21]***
Year == 2000	1681.39	1971.78	1648.71	1,595.50
	[3.95]***	[4.50]***	[3.78]***	[3.66]***
Year == 2001	2448.39	3228.46	2620.66	2,452.62
	[4.00]***	[4.71]***	[4.29]***	[4.23]***
Year == 2002	2867.33	4135.11	3268.36	3,000.18
	[4.25]***	[5.24]***	[4.61]***	[4.44]***
Year == 2003	2409.60	4439.87	3390.61	2,989.08
	[3.43]***	[5.26]***	[4.32]***	[3.63]***
Year == 2004	248.42	3083.74	2058.11	1,514.71
	[0.30]	[4.00]***	[2.53]**	[1.50]
Year == 2005	395.26	3928.77	2760.89	2,076.97
	[0.42]	[4.63]***	[3.08]***	[1.66]*
Year == 2006	1841.05	5560.91	4201.87	3,479.14
	[1.73]*	[5.25]***	[3.92]***	[2.47]**
Federal Funding, Other Fields			0.08	0.06
			[1.28]	[0.86]
Non-Federal Funding, Other Fields			0.30	0.29
			[2.12]**	[2.05]**
Observations	2176	2176	2176	2176
Number of institutions	272	272	272	272

* significant at 10%; ** significant at 5%; *** significant at 1%

Robust t statistics in brackets

Results from multivariate regression with non-federal life sciences funding as the dependent variable, and with the university-year as unit of observation. All federal funding amounts are lagged one year. Model 1 includes year fixed effects (shown); Model 2 adds university fixed effects. Model 3 adds controls for non-life-sciences funding received by the university. Model 4 uses lagged predicted NIH funding as an instrument for federal life sciences funding, while also including year fixed effects and controls for non-life-sciences funding received by the university.

3.2 Effects of University Heterogeneity

In this section, we investigate how effects of federal funding differ by the following university characteristics: (1) tercile of non-federal life sciences R&D funding received in the base year (1998), (2) tercile of federal life sciences R&D funding received in the base year (1998), (3) PhD versus non-PhD-granting status, (4) private versus public institutional control, and (5) whether the institution was ranked among the top 50 research universities by U.S. News & World Report. All forty of the institutions in our panel that ranked among the top 50 in U.S. News & World Report at some point during our sample period are PhD-granting institutions, 25 are private, 27 are in the top tercile for non-federal life sciences R&D funding, and 33 are in the top tercile for federal life sciences R&D funding.

Table 3 columns (2) through (4) show results of IV regressions by tercile of non-federal life sciences funding in 1998, the first year of our panel. Federal life sciences R&D funding leads to a statistically significant increase in non-federal funding for universities in all three terciles. The greatest effect appears to be in the lowest tercile: for universities that historically have received relatively low non-federal R&D funding for the life sciences, a dollar increase in federal funding yields a $0.53 increase in non-federal funding, compared with $0.16 and $0.23 for the second and third terciles.

Columns (5) through (7) report results of IV regressions by tercile of federal life sciences funding in 1998. We again find federal life sciences R&D funding is associated with a statistically significant increase ($p<.10$) in non-federal funding for universities in all three terciles, and again the greatest effect seems to be in the lowest tercile: for universities that historically received relatively low federal R&D funding for the life sciences, a dollar increase in

federal funding yields an estimated $0.91 increase in non-federal funding, compared with $0.54 and $0.21 for the higher two terciles.

Table 4 columns (2) and (3) show results by highest degree granted. The effects of federal funding are much larger for universities granting only Masters or Bachelors degrees versus those awarding PhDs in S&E fields. While a dollar increase in federal life sciences R&D funding at a PhD-granting institution is associated with a $0.32 increase in non-federal funding, for institutions that do not grant a PhD, a dollar increase in federal funding yields a $1.10 increase in non-federal funding. This difference is statistically significant (α=.05).

In columns (4) and (5), we report results for public versus private universities. At public universities, a dollar increase in federal life sciences R&D funding is associated with a $0.41 increase in non-federal funding, whereas at private universities a dollar increase in federal life sciences R&D funding is associated with a $0.25 increase in non-federal funding.

Finally, in columns (6) and (7), we find the non-federal funding response seems also to be greater for universities not among the top 50 in the U.S. News and World Report rankings. For these universities, a dollar increase in federal life sciences R&D funding yields only a $0.19 increase in non-federal funding. For all other universities (including top-ranked liberal arts colleges), the effect appears to be higher: a dollar increase in federal life sciences R&D funding yields a $0.46 increase in non-federal funding.

Table 3: Heterogeneous effects of federal funding by base-year levels of federal and non-federal funding

	Terciles of Non-Federal Funding			Terciles of Federal Funding		
	Tercile 1	Tercile 2	Tercile 3	Tercile 1	Tercile 2	Tercile 3
Federal Funding for Life Sciences R&D	**0.53**	**0.16**	**0.23**	**0.91**	**0.54**	**0.21**
	[1.85]*	[3.77]***	[1.87]*	[1.86]*	[3.33]***	[1.48]
Federal R&D Funding, Other Fields	-0.03	-0.02	0.05	-0.09	-0.01	0.08
	[0.54]	[0.57]	[0.53]	[0.92]	[0.33]	[0.85]
Non-Federal R&D Funding, Other Fields	0.03	0.02	0.58	0.03	0.09	0.32
	[2.70]***	[0.54]	[3.40]***	[1.83]*	[2.21]**	[1.70]*
Year==2000	136.76	602.90	4,380.07	152.52	263.86	5,203.58
	[1.11]	[3.20]***	[3.06]***	[1.22]	[0.78]	[3.68]***
Year==2001	-75.29	1,108.47	7,506.69	-40.91	721.86	8,608.93
	[0.84]	[3.92]***	[3.99]***	[0.60]	[1.54]	[4.09]***
Year==2002	-51.76	1,284.63	9,569.95	-51.06	803.94	11,082.10
	[0.55]	[3.82]***	[3.84]***	[0.45]	[1.60]	[3.89]***
Year==2003	-15.05	1,946.86	10,501.97	-35.66	580.87	12,562.27
	[0.27]	[4.60]***	[2.88]***	[0.52]	[0.99]	[2.97]***
Year==2004	-54.39	1,785.99	8,593.12	-34.32	197.10	9,945.27
	[1.04]	[3.50]***	[1.70]*	[0.43]	[0.25]	[1.67]*
Year==2005	15.52	2,485.34	11,549.68	-20.89	-228.73	12,883.80
	[0.23]	[4.66]***	[1.90]*	[0.27]	[0.24]	[1.76]*
Year==2006	-52.29	3,267.04	15,484.97	-75.82	279.24	17,163.11
	[0.59]	[5.29]***	[2.34]**	[0.89]	[0.24]	[2.14]**
Observations	720	736	720	720	736	720
Number of Institutions	90	92	90	90	92	90

* significant at 10%; ** significant at 5%; *** significant at 1%

Robust t-statistics in brackets below each coefficient estimate.

Results from instrumental variable regressions with non-federal life sciences funding as the dependent variable, and with the university-year as unit of observation. All federal funding amounts are lagged one year. All models include year fixed effects (shown) and university fixed effects. Standard errors are robust to heteroskedasticity and clustered on university.

Table 4: Heterogeneous effects of federal funding by highest degree granted, institutional control, and U.S. News & World Report ranking

	Highest Degree Granted		Control		Rank	
	PhD	Other	Private	Public	Top 50	Other
Federal Funding for Life Sciences R&D	**0.32**	**1.05**	**0.25**	**0.41**	**0.19**	**0.46**
	[3.87]***	[11.33]***	[5.83]***	[3.04]***	[1.77]*	[3.69]***
Federal R&D Funding, Other Fields	0.07	-0.04	0.05	0.00	0.09	0.08
	[0.94]	[0.41]	[0.76]	[0.02]	[0.98]	[0.83]
Non-Federal R&D Funding, Other Fields	0.30	0.07	0.24	0.30	0.68	0.17
	[2.02]**	[1.12]	[1.65]	[1.81]*	[2.59]**	[1.38]
Year==2000	2,079.57	-249.65	358.82	2,172.12	5,341.85	930.71
	[3.74]***	[0.62]	[1.00]	[3.52]***	[2.86]***	[2.32]**
Year==2001	3,231.64	-500.67	855.82	3,163.80	6,379.31	1,544.15
	[4.37]***	[0.98]	[1.30]	[3.92]***	[2.40]**	[3.50]***
Year==2002	3,868.08	-133.41	818.62	3,984.31	7,474.08	1,865.90
	[4.44]***	[0.30]	[1.25]	[4.09]***	[2.30]**	[3.35]***
Year==2003	3,791.71	-48.87	1,023.60	3,809.91	7,368.01	1,585.03
	[3.44]***	[0.18]	[1.37]	[3.14]***	[1.89]*	[2.20]**
Year==2004	2,045.65	-1,035.44	-1,017.87	2,548.96	2,776.16	631.40
	[1.47]	[2.29]**	[1.01]	[1.75]*	[0.54]	[0.63]
Year==2005	2,763.35	-1,058.84	914.72	2,360.31	7,286.66	241.56
	[1.60]	[2.25]**	[1.21]	[1.22]	[1.18]	[0.18]
Year==2006	4,426.00	-303.23	1,801.19	3,966.84	8,629.00	1,819.57
	[2.28]**	[0.91]	[2.20]**	[1.83]*	[1.21]	[1.37]
Observations	1,696.00	480	688	1488	320	1856
Number of Institutions	212	60	86	186	40	232

* significant at 10%; ** significant at 5%; *** significant at 1%

Robust t-statistics in brackets below each coefficient estimate.

Results from instrumental variable regressions with non-federal life sciences funding as the dependent variable, and with the university-year as unit of observation. All federal funding amounts are lagged one year. All models include year fixed effects (shown) and university fixed effects. Standard errors are robust to heteroskedasticity and clustered on university.

In addition to these analyses, we also investigated whether contemporaneous and lagged federal funding up to three years prior had significant impact on non-federal funding in the life sciences. We chose a maximum lag of three years, because the average noncompeting duration for NIH awards is four years. In a finite distributed lag model with university fixed effects, we find no evidence of substitution, i.e., none of the coefficients for federal funding are significant and negative. Per Wooldridge (2009), the sum of the coefficients on contemporaneous and lagged federal R&D can be understood as the effect of a sustained increase in federal life sciences R&D funding to a university. We find that a sustained one dollar increase in federal funding is associated with a $0.25 increase in non-federal funding for life sciences R&D (p=.0001).

3.3 Validation of the IV model

The above IV model will produce unbiased causal estimates of the impact of federal life sciences funding on non-federal life sciences funding, as long as our instrument satisfies two key assumptions: first, predicted NIH funding must be strongly correlated with actual federal life sciences funding; and second, predicted NIH funding must be uncorrelated with any unobserved time-varying university characteristics that impact non-federal life sciences funding. The first assumption is testable, and we do so by computing the first-stage F-statistic for our instrument. However, the second assumption cannot be tested directly. As noted in equation (4), variation in our instrument over time is based on the share of a university's funding received from a particular NIC in 1997 and differential growth in aggregate NIC budgets in the following years. University fixed effects control for differences in total NIH funding across universities in 1997, so the variation in our instrument arises only from year-to-year changes in the NIC budgets. In essence, we predict higher NIH funding over time for universities that, in 1997, specialized in

research areas funded by NICs which, in turn, experienced greater growth from 1998 through 2006. The validity of our second assumption relies on the following notions: (a) any shock to a given NIC's budget for a given year is uncorrelated with the amount of non-federal life sciences R&D funding a university would receive the following year, except to the extent that non-federal life sciences funding is itself dependent on NIH funding, and (b) university specialization in the base year is uncorrelated with other institutional characteristics that make the university more or less likely to obtain non-federal funding in later periods (e.g., institutional reputation). We believe these are plausible assumptions, for the reasons discussed below.

3.3.1 Determination of budgets for the National Institutes of Health is unlikely to be related to non-federal research priorities

The National Institutes of Health provide extramural research funding to universities in all fifty states. Each NIC specializes in specific diseases, aspects of human health and development, or research support, and each is funded by a separate Congressional appropriation. The U.S. Senate and House of Representatives have established parallel appropriations subcommittees that jointly determine the budget for the NICs, along with competing budgetary priorities in other health and human services agencies (e.g., the CDC), education, and labor.

The appropriations cycle begins with the President's submission of a recommended budget to Congress. The President's recommendations are accompanied by detailed justification statements from each federal agency, including from each of the NICs. Individual members of Congress then submit prioritized "wish lists" to the appropriations subcommittee based on requests from their constituents, including disease advocacy and general science lobby groups, academic institutions, etc. (Kennan, 2005). Both Congressional subcommittees hold hearings with testimony from NIC officials. Once the subcommittees receive their spending ceilings, the House committee traditionally acts first to "mark up" an appropriation bill, and the Senate passes

an amended version (for further discussion, see Streeter (2006)). Thus, decisions on the portion of the budget designated for each NIC are made in response to political demands and perceived unmet public health needs (e.g., for bioterrorism-related research after 2001), but in tandem with funding decisions for unrelated aspects of labor and education due to their common spending cap. After the House and Senate have passed their individual versions of the appropriations bill, differences are resolved in conference, resulting in a final product called the conference report. This conference report provides budget appropriations for each NIC, and may include additional non-binding "report language" encouraging the NICs to pursue research in particular areas or provide particular funding mechanisms. Within each NIC, however, Congress has generally avoided specifying amounts for particular fields of research or funding mechanisms, and the number of such directives has also declined in recent years (Committee on the NIH Research Priority-Setting Process and Institute of Medicine, 1998). For example, the 2009 Omnibus Appropriations Bill directs the National Cancer Institute only in the naming of a surgical fellowship.

Research priorities within each NIC (and, therefore, the amount of funding available to universities specializing in one research field versus another) are determined by the NIC in consultation with the Office of the Director. Decision-making with respect to funding particular proposals is highly decentralized, and funding levels by disease or specialized research area reflect not only the social and economic costs of particular diseases, but also the quality of investigator-initiated proposals received by each NIC and their potential for scientific progress. Thus, while it is certainly conceivable that technological opportunity or some other root cause might simultaneously increase both federal and non-federal funding for a particular disease (and thus potentially provide increased availability of both types of funding for a particular

university), such within-NIC research priority shifts are unlikely to be reflected in the aggregate appropriation by Congress to a particular NIC.

As a first test of this potential confounder, we generated an alternate measure of predicted NIH funding for university *u* in year *t*, based solely on each university's share of total NIH funding in the base year, and changes in total (aggregate) NIH funding over time. Our results with this alternate instrument are essentially identical to the original IV approach: we find a dollar increase in federal life sciences funding results in a $0.36 increase in non-federal life sciences funding.

Next, we investigated whether increases in NIH funding by institute were associated with contemporaneous increases in industry R&D for the diseases represented by each NIC. The dependent variable for this analysis is industry R&D as measured by the number of drugs entering clinical trials for each disease category (matched to the relevant NIC, as presented in the Appendix) and year. Matching was based primarily on the lead NIC identified for each disease, as reported by the MedlinePlus website at the National Library of Medicine; however, this assignment was supplemented by analysis of the fraction of grants awarded by each NIC that were classified to each disease in 2006. The algorithm for this classification is detailed in Blume-Kohout (2009). For example, whereas the primary NIH organization for research on allergies is the National Institute of Allergy and Infectious Diseases (NIAID), we found extramural research grants for *respiratory* allergies were more commonly funded by the National Heart, Lung, and Blood Institute (NHLBI), and so assigned treatments for respiratory allergies to that NIC. The key explanatory variable is *log*(NIC Budget) where the dollar amounts are adjusted by the Biomedical Research and Development Price Index (BRDPI).

Due to evidence of moderate overdispersion, in lieu of Poisson we estimated an unconditional negative binomial model with NIC and year fixed effects. We found no significant relationship between NIH funding and contemporaneous pharmaceutical R&D (coefficient -.082, p=.833, with robust standard errors clustered on NIC). Expressing NIC budget funding in levels rather than logs yielded very similar results. These results support our assumption that shocks to universities' federal R&D funding due to changes in NIC budgets are unlikely to be accompanied by contemporaneous shocks to university funding from industry sources.

3.3.2 University baseline shares are uncorrelated with other characteristics affecting likelihood of receiving non-federal funding

The university baseline shares in 1997 used for our instrument are intended to indicate the specialization of each university across research fields, and research fields are strongly tied to particular NICs. If universities with stronger research reputations, or other characteristics affecting their likelihood of receiving non-federal funding, specialized in research areas that were funded by the NICs that grew most rapidly after 1997, this could invalidate our instrument. To investigate this possibility, we tested whether the share of funding each university received from each NIC in the base year was correlated with other observable characteristics found to influence non-federal funding, including whether the university grants PhD degrees, whether the university is private or public, whether the university is ranked among the top 50 research universities in U.S. News and World Report, and how much funding the university received in fields other than life sciences during the sample period. To implement this test, for each NIC we regressed the share of each university's NIH funding in 1997 awarded by that NIC on these university characteristics. Applying the Bonferroni correction α/N for multiple regressions, where $\alpha = .05$ and N represents the 22 NICs that funded universities in our panel in the base year, we find only one significant association between university characteristics and NIC funding

share: universities granting PhDs received disproportionately more funding from the National Cancer Institute in our base year than universities granting only bachelor's and master's degrees. Public versus private, top 50 versus unranked, and university non-life-sciences funding levels (both federal and non-federal) all had no impact on the share of funding universities received from each NIC. For the 21 other NICs, PhD-granting also had no significant impact.

Finally, we tested whether our results are robust to using an earlier base year for our analysis. Doing so guards against the possibility that, based on information available about future funding changes, universities had already altered their specialization in the benchmark base year of 1997. Such anticipatory behavior is unlikely for a base year further removed from the sample period. To implement this test, we calculate NIC shares for each university in 1992 instead of 1997, predict NIH funding for each university, and use this alternate measure of predicted NIH funding to instrument for federal life science funding. The results are almost identical to our main analysis: for example, we estimate $0.36 in non-federal funding per federal life sciences dollar ($p<.001$), with point elasticity (at the average) of 0.59.

3.3.3 IV results show very little correlation between federal life sciences funding and non-federal funding in seemingly unrelated fields

As an additional validity check, we use our instrumental variables approach to test whether changes in federal life sciences funding were associated with changes in non-federal funding for other S&E fields including social sciences, psychology, computer science, engineering, physical sciences, mathematical sciences, and environmental sciences. In principle, we should expect little relationship between federal life sciences funding and non-federal funding in these largely unrelated fields, provided our IV approach controls for differences in fundraising effort, university reputation, and other time-varying university characteristics. To the extent there is spillover, we would expect it mainly in the other, non-life-sciences fields that

NIH also funds. In 2006, approximately 84% of NIH funding for basic and applied research at U.S. universities was for life sciences, 6% was for psychology, 3% was for engineering, 1.7% was for physical sciences, 1.4% was for environmental sciences, 1% was for social sciences, and less than 1% was for mathematical and computer sciences.[6] As expected, we find little causal relationship between federal life science funding (using the predicted NIH funding instrument) and non-federal funding in these fields, The only significant effects found were for non-federal social sciences funding (1.8 cents per federal dollar), physical sciences (1.3 cents per federal dollar), environmental sciences (0.7 cents per federal dollar), and mathematics (0.2 cents per federal dollar), generally tracking the NIH funding levels described above, and much lower than the 33-cent increase seen for non-federal life sciences funding. The sole exception to this pattern was for social sciences research, which may simply reflect differences in overall availability of nonfederal funding for social sciences versus other fields.

3.3.4 Inclusion of time-varying university rankings has no impact on our results

If our IV approach were invalid, then changes in the quality or reputation of a university over time would likely impact both federal and non-federal funding, confounding our results. To test for this possibility, we include as a covariate the U.S. News and World Report rank, year-by-year, for each university in our panel that at some point during the sample period was ranked among the top 50 national universities.[7] Changes in university rank over time had no impact on our estimate, and the coefficient on rank was insignificant (robust t-statistic 0.38), suggesting that university fixed effects coupled with our IV approach are adequate to describe the impact of university ranking on R&D funding.

4. Discussion

As discussed above, the effect of public (federal) research funding on private or non-federal funding is an issue of much debate. A priori, there could be a negative (substitution) effect, either due to crowding out of private investment or because researchers could stop seeking other sources of funding once they receive federal funding. On the other hand, federal R&D funding could also have a positive impact due to complementarity or signaling effects. Careful empirical analysis is needed to parse these opposing effects.

In a panel of 272 U.S. universities, with data on federal and non-federal life sciences R&D funding spanning nearly a decade, we find that increased federal funding is associated with increased non-federal funding. This result is robust to econometric specifications that correct for omitted latent variables that could be responsible for increases in both types of funding. Our evidence also suggests that, in addition to providing support for research, successful applications for federal life sciences funding may be interpreted by private and non-federal organizations as a signal of the quality of the recipient institutions. Less research-focused universities (i.e., those not granting PhDs in S&E fields), and those that historically received lower levels of federal and non-federal funding appear to experience greater increases in non-federal funding for each federal dollar received. This is consistent with the observation by Payne (2001) that federal organizations such as the NIH may serve to correct asymmetries in information between universities and their prospective non-federal funders.

Our result is also qualitatively similar to that found by Diamond (1999), who used aggregate, annual time series data from 1953-1995 to study the connection between federal and private spending on basic research in science. However, his estimate is quantitatively much lower than ours: he found that a dollar increase in federal spending was associated with only an

additional $0.08 in academic spending, compared with $0.33 in our analysis. Our use of the university as a unit of analysis, and reliance on cross-sectional as well as time-series variation (as opposed to aggregate time-series variation alone) might partly account for this difference.

By enabling universities to attract more private resources, federal life sciences R&D funding may also influence the eventual commercialization of university research. However, to understand the mechanisms by which federal funding results in commercial products such as life-saving drugs, one should look at broader outcomes such as university patenting and licensing behavior, and alliances between universities and the private sector. Qualitative research with not-for-profit foundations and other non-federal funders could serve to confirm the suggested signaling effect; however, a more structural approach is needed to disentangle this effect from complementarity between federal and non-federal funds in the production of knowledge. These are subjects for our future research.

Notes

[1] Federal obligations for biological, medical, and life sciences not elsewhere classified totaled 83.5% of NIH funding to universities and colleges. Data are based on author's calculations using the NSF Survey of Federal Funds for Research and Development, available at: http://caspar.nsf.gov/

[2] We are very grateful to Bhaven Sampat at Columbia University who provided us with analytic data files extracted from NIH administrative records for years 1997-2003, including the financial amount of each award, and to Pierre Azoulay who provided a list of NIH grantee organizations coded by type (e.g., institutions of higher education). These data were obtained via FOIA requests. Data for years 2004-2006 were downloaded from the NIH website at: http://www.report.nih.gov/award/awardtr.cfm

[3] The NSF estimated total life sciences R&D expenditures at U.S. universities and colleges as $28.8 billion in 2006. See: http://www.nsf.gov/statistics/nsf08300/pdf/tab4.pdf

[4] While the linear model does exhibit increasing variance in the residuals with higher levels of non-federal funding, there is no evidence of any upward or downward trend suggesting misspecification, and log transformation of the financial amounts appears to overcorrect rather than stabilize the variance. Thus, the data do not recommend log transformation, our estimates should be both unbiased and consistent except as discussed above, and the observed heteroskedasticity can be accommodated simply by our calculation of robust standard errors. We calculate robust clustered standard errors for the instrumental variables models using XTIVREG2 (Schaffer, 2005).

[5] Despite our misgivings about this alternate specification, we did nonetheless estimate each of our models described above for this restricted panel, using log-transformed financial amounts.

For models (1)-(3), we find significant positive coefficients ($p<.01$) ranging from 0.925 for model (1) to 0.492 for model (3). However, for this restricted panel of institutions already receiving non-zero federal and non-federal funding both for life sciences and for other S&E fields, the instrumental variables approach with constant-elasticity specification indicates no significant impact of lagged federal life sciences funding on non-federal life sciences funding. Rather, for these institutions, receipt of non-federal life sciences funding appears to be most strongly and significantly correlated with the university's contemporaneous successes in attracting non-federal funding for other science and engineering fields.

[6] Estimates for NIH obligations to U.S. universities for basic and applied research, by field, are taken from the NSF's Survey of Federal Funds for Research and Development.

[7] Historical rankings were obtained from the Chronicle of Higher Education website, accessed April 23, 2009, at: http://chronicle.com/stats/usnews. Forty universities in our panel had ranking data available for at least 8 out of 9 panel-years.

References

Blume-Kohout, Margaret. "Federally-funded biomedical research and pharmaceutical research and development." In *Essays on Government Policy and Pharmaceutical Innovation*. Santa Monica, CA: Pardee RAND Graduate School, 2009.

Committee on the NIH Research Priority-Setting Process and Institute of Medicine. "Scientific opportunities and public needs: improving priority setting and public input at the National Institutes of Health." Washington DC: National Academies Press, 1998.

David, Paul A., Bronwyn H. Hall, and Andrew A. Toole, "Is public R&D a complement or substitute for private R&D? A review of the econometric evidence," *Research Policy,* Vol. 29, 2000.

Diamond, Arthur M., "Does federal funding 'crowd in' private funding of science?," *Contemporary Economic Policy,* Vol. 17, No. 4, 1999, pp. 423-431.

Jacob, Brian and Lars Lefgren, *The impact of research grant funding on scientific productivity*, NBER Working Paper 13519, 2007.

Kennan, Stephanie A. "Legislative relations in public health." In *Public Health Administration*, Novick, Lloyd F. and Glen P. Mays, eds. Boston: Jones & Bartlett, 2005.

Lawler, Andrew, "University-industry collaboration: Last of the big-time spenders?," *Science,* Vol. 299, No. 5605, 2003, pp. 330-333.

National Institutes of Health, "Summary of the FY 2009 President's Budget," 2008, http://officeofbudget.od.nih.gov/pdfs/FY09/Summary%20of%20FY%202009%20Budget-Press%20Release.pdf (as of April 14).

Payne, A. Abigail, "Measuring the effect of federal research funding on private donations at research universities: is federal research funding more than a substitute for private donations?," *International Tax and Public Finance,* Vol. 8, 2001, pp. 731-751.

Payne, A. Abigail and Aloysius Siow, "Does federal research funding increase university research output?," *Advances in Economic Analysis & Policy,* Vol. 3, No. 1, 2003.

Schaffer, Mark, XTIVREG2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models, last revised 16 Oct 2008, Boston College Department of Economics, 2005.

Smith, Pamela W., *The National Institutes of Health (NIH): Organization, Funding, and Congressional Issues*, Washington, DC: Congressional Research Service, October 19, 2006.

Streeter, Sandy, *The Congressional appropriations process: an introduction*, Washington, DC: Congressional Research Service, 2006.

Wooldridge, Jeffrey M., *Introductory Econometrics: A Modern Approach*, 4th ed. ed., Mason, OH: South-Western Cengage Learning, 2009.

Appendix. Crosswalk between Pharmaprojects therapeutic codes and primary NIH Institutes, page 1 of 5

Pharmaprojects	Therapeutic Category	NIH Institute
A10B	Antidiabetic	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
A10C	Symptomatic antidiabetic	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
A11A	Nutritional supplement	Not Assigned
A14	Anabolic	National Institute of Neurological Disorders and Stroke (NINDS)
A15	Appetite stimulant	National Institute of Neurological Disorders and Stroke (NINDS)
A16	GI inflammatory/bowel disorders	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
A17	Metabolic and enzyme disorders	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
A18Z	Alimentary/Metabolic, other	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
A1A	Stomatological	National Institute of Dental and Craniofacial Research (NIDCR)
A2A	Antacid/Antiflatulent	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
A2B	Antiulcer	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
A3	Antispasmodic	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
A4A	Antiemetic	National Cancer Institute (NCI)
A4B	Gastroprokinetic	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
A5B	Hepatoprotective	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
A5D	Gallstone therapy	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
A6	Laxative	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
A7	Antidiarrhoeal	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
A8A3	Anorectic/Antiobesity	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
A9	Digestive	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
B1A	Anticoagulant	National Heart, Lung, and Blood Institute (NHLBI)
B1B1	Fibrinolytic	National Heart, Lung, and Blood Institute (NHLBI)
B1B9	Antithrombotic	National Heart, Lung, and Blood Institute (NHLBI)
B2A	Antifibrinolytic	National Heart, Lung, and Blood Institute (NHLBI)
B2B	Haemostatic	National Heart, Lung, and Blood Institute (NHLBI)
B3C1	Antisickling	National Heart, Lung, and Blood Institute (NHLBI)
B3C9	Antianaemic	National Heart, Lung, and Blood Institute (NHLBI)
B5A1	Plasma substitute	National Heart, Lung, and Blood Institute (NHLBI)
B5A2	Blood fraction	National Heart, Lung, and Blood Institute (NHLBI)
B6A	Septic shock treatment	National lnstitute of General Medical Sciences (NIGMS)
B7Z	Haematological	National Heart, Lung, and Blood Institute (NHLBI)

Appendix. Crosswalk between Pharmaprojects therapeutic codes and primary NIH Institutes, page 2 of 5

Pharmaprojects	Therapeutic Category	NIH Institute
C10	Hypolipaemic/Antiatherosclerosis	National Heart, Lung, and Blood Institute (NHLBI)
C1B	Antiarrhythmic	National Heart, Lung, and Blood Institute (NHLBI)
C1C1	Cardiostimulant	National Heart, Lung, and Blood Institute (NHLBI)
C1D1	Vasodilator, coronary	National Heart, Lung, and Blood Institute (NHLBI)
C1D3	Antianginal	National Heart, Lung, and Blood Institute (NHLBI)
C2B1	Antihypertensive, adrenergic	National Heart, Lung, and Blood Institute (NHLBI)
C2B2	Antihypertensive, renin system	National Heart, Lung, and Blood Institute (NHLBI)
C2B6	Antihypertensive, diuretic	National Heart, Lung, and Blood Institute (NHLBI)
C2B9	Antihypertensive, other	National Heart, Lung, and Blood Institute (NHLBI)
C4A	Vasodilator, peripheral	National Heart, Lung, and Blood Institute (NHLBI)
C4B	Vasodilator, renal	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
C5A	Vasoprotective, topical	National Institute on Aging (NIA)
C5C	Vasoprotective, systemic	National Heart, Lung, and Blood Institute (NHLBI)
C6C	Hypertensive	National Heart, Lung, and Blood Institute (NHLBI)
C9Z	Cardiovascular	National Heart, Lung, and Blood Institute (NHLBI)
D10A	Antiacne	National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
D11Z	Dermatological	National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
D3A	Vulnerary	National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
D4A	Antipruritic/inflamm, allergic	National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
D4B	Antipruritic/inflamm, non-allergic	National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
D5A	Antipsoriasis	National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
G1C	Fertility enhancer	National Institute of Child Health & Human Development (NICHD)
G2A	Labour inducer	National Institute of Child Health & Human Development (NICHD)
G2B	Labour inhibitor	National Institute of Child Health & Human Development (NICHD)
G3A	Menstruation disorders	National Institute of Child Health & Human Development (NICHD)
G3B	Menopausal disorders	National Institute on Aging (NIA)
G3C	Female contraceptive	National Institute of Child Health & Human Development (NICHD)
G3D	Abortifacient	National Institute of Child Health & Human Development (NICHD)
G4Z	Urological	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
G5A	Prostate disorders	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
G5B	Male sexual dysfunction	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)

Appendix. Crosswalk between Pharmaprojects therapeutic codes and primary NIH Institutes, page 3 of 5

Pharmaprojects	Therapeutic Category	NIH Institute
G5C	Male contraceptive	National Institute of Child Health & Human Development (NICHD)
G6Z	Reproductive/gonadal, general	National Institute of Child Health & Human Development (NICHD)
H1A	ACTH	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
H3A	Thyroid hormone	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
H3B	Antithyroid	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
H4B	Prostaglandin	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
H4C	Releasing hormone	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
H4D2	Antiprolactin	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
H4E1	Insulin	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
H4E2	Glucagon	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
H4F1	Growth hormone	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
H4F2	Somatostatin	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
H4Z	Hormone	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
I1A	Immunostimulant, anti-AIDS	National Institute of Allergy and Infectious Diseases (NIAID)
I1Z	Immunostimulant, other	National Institute of Allergy and Infectious Diseases (NIAID)
I2	Cytokine	National Institute of Allergy and Infectious Diseases (NIAID)
I4A2	Immunoglobulin, non-MAb	National Institute of Allergy and Infectious Diseases (NIAID)
I5	Immunosuppressant	National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
I6Z	Immunological	Not Assigned
J1A	Tetracycline	National Institute of Allergy and Infectious Diseases (NIAID)
J1C1	Penicillin, oral	National Institute of Allergy and Infectious Diseases (NIAID)
J1C2	Penicillin, injectable	National Institute of Allergy and Infectious Diseases (NIAID)
J1D1	Cephalosporin, oral	National Institute of Allergy and Infectious Diseases (NIAID)
J1D2	Cephalosporin, injectable	National Institute of Allergy and Infectious Diseases (NIAID)
J1E	Trimethoprim and analogues	National Institute of Allergy and Infectious Diseases (NIAID)
J1F	Macrolide antibiotic	National Institute of Allergy and Infectious Diseases (NIAID)
J1L	Aminoglycoside antibiotic	National Institute of Allergy and Infectious Diseases (NIAID)
J1M	Peptide antibiotic	National Institute of Allergy and Infectious Diseases (NIAID)
J1N	Beta-lactam antibiotic	National Institute of Allergy and Infectious Diseases (NIAID)
J1Z	Antibiotic, other	National Institute of Allergy and Infectious Diseases (NIAID)
J2A	Antifungal	National Institute of Allergy and Infectious Diseases (NIAID)

Appendix. Crosswalk between Pharmaprojects therapeutic codes and primary NIH Institutes, page 4 of 5

Pharmaprojects	Therapeutic Category	NIH Institute
J3C	Quinolone antibacterial	National Institute of Allergy and Infectious Diseases (NIAID)
J3Z	Antibacterial, other	National Institute of Allergy and Infectious Diseases (NIAID)
J4A	Antimycobacterial	National Institute of Allergy and Infectious Diseases (NIAID)
J5A	Antiviral, anti-HIV	National Institute of Allergy and Infectious Diseases (NIAID)
J5B	Antiviral, interferon	National Institute of Allergy and Infectious Diseases (NIAID)
J5Z	Antiviral, other	National Institute of Allergy and Infectious Diseases (NIAID)
J7A1	Prophylactic vaccine	National Institute of Allergy and Infectious Diseases (NIAID)
J7A2	Therapeutic vaccine	National Institute of Allergy and Infectious Diseases (NIAID)
J7B	Immunomodulator, anti-infective	National Institute of Allergy and Infectious Diseases (NIAID)
J8Z	Anti-infective, other	National Institute of Allergy and Infectious Diseases (NIAID)
K1A	Anticancer, antibiotic	National Cancer Institute (NCI)
K1B	Anticancer, alkylating	National Cancer Institute (NCI)
K1C	Anticancer, antimetabolite	National Cancer Institute (NCI)
K2	Anticancer, hormonal	National Cancer Institute (NCI)
K3	Anticancer, immunological	National Cancer Institute (NCI)
K4	Anticancer, interferon	National Cancer Institute (NCI)
K5A	Radio/chemosensitizer	National Cancer Institute (NCI)
K5B	Radio/chemoprotective	National Cancer Institute (NCI)
K6Z	Anticancer, other	National Cancer Institute (NCI)
M1A1	Anti-inflammatory	National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
M1A2	Anti-inflammatory, topical	National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
M2C	Antiarthritic, immunological	National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
M2Z	Antiarthritic, other	National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
M3	Muscle relaxant	National Institute of Neurological Disorders and Stroke (NINDS)
M4A	Antigout	National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
M5A	Osteoporosis treatment	National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
M5Z	Musculoskeletal	National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
N10A	Antidepressant	National Institute of Mental Health (NIMH)
N11Z	Neurological	National Institute of Mental Health (NIMH)
N1A1	Anaesthetic, inhalation	National Institute of General Medical Sciences (NIGMS)
N1A2	Anaesthetic, injectable	National Institute of General Medical Sciences (NIGMS)

Appendix. Crosswalk between Pharmaprojects therapeutic codes and primary NIH Institutes, page 5 of 5

Pharmaprojects	Therapeutic Category	NIH Institute
N1B	Anaesthetic, local	National Institute of General Medical Sciences (NIGMS)
N2B	Analgesic, NSAID	National Institute of Neurological Disorders and Stroke (NINDS)
N2Z	Analgesic, other	National Institute of Neurological Disorders and Stroke (NINDS)
N3A	Antiepileptic	National Institute of Neurological Disorders and Stroke (NINDS)
N4A	Antiparkinsonian	National Institute of Neurological Disorders and Stroke (NINDS)
N5A1	Neuroleptic	National Institute of Mental Health (NIMH)
N5B	Hypnotic/Sedative	National Institute of Mental Health (NIMH)
N5C	Anxiolytic	National Institute of Mental Health (NIMH)
N5D	Antineurotic	National Institute of Mental Health (NIMH)
N6B	Psychostimulant	National Institute of Mental Health (NIMH)
N6D	Cognition enhancer	National Institute on Aging (NIA)
N7A	Multiple sclerosis treatment	National Institute of Neurological Disorders and Stroke (NINDS)
N7C	Neuroprotective	National Institute of Neurological Disorders and Stroke (NINDS)
N8A	Antimigraine	National Institute of Neurological Disorders and Stroke (NINDS)
N9A	Dependence treatment	National Institute on Drug Abuse (NIDA)
P1A	Amoebicide	National Eye Institute (NEI)
P1B	Anthelmintic	National Institute of Allergy and Infectious Diseases (NIAID)
P1C	Schistosomicide	National Institute of Allergy and Infectious Diseases (NIAID)
P1D	Antimalarial	National Institute of Allergy and Infectious Diseases (NIAID)
P1G	Protozoacide	National Institute of Allergy and Infectious Diseases (NIAID)
P1Z	Parasiticide	National Institute of Allergy and Infectious Diseases (NIAID)
R3A	Lung Surfactant	National Heart, Lung, and Blood Institute (NHLBI)
R4A	COPD treatment	National Heart, Lung, and Blood Institute (NHLBI)
R4B	Cystic fibrosis treatment	National Heart, Lung, and Blood Institute (NHLBI)
R5D	Antitussive	National Heart, Lung, and Blood Institute (NHLBI)
R8A	Antiasthma	National Heart, Lung, and Blood Institute (NHLBI)
R8B	Antiallergic, non-asthma	National Heart, Lung, and Blood Institute (NHLBI)
R9A	Respiratory stimulant	National Heart, Lung, and Blood Institute (NHLBI)
R9Z	Respiratory	National Heart, Lung, and Blood Institute (NHLBI)
S1G	Antiglaucoma	National Eye Institute (NEI)
S1Z	Ophthalmological	National Eye Institute (NEI)
S2	Otological	National Institute on Deafness and Other Communication Disorders (NIDCD)